AD-A093 543 WHITE SANDS MISSILE RANGE NM F/G 12/1
ROBUST FILTERING AND SMOOTHING VIA GAUSSIAN MIXTURES.(U)
DEC 80 W S AGEE, B A DUNN
UNCLASSIFIED WSMR-TR-73 NL

END
DATE
FILMED
2 81
DTIC

(12)

AD A093543

TECHNICAL REPORT

NO. 73

JANUARY 1980

LEVEL II

ROBUST FILTERING AND SMOOTHING VIA GAUSSIAN MIXTURES

DTIC

MATHEMATICAL SERVICES BRANCH
DATA SCIENCES DIVISION
US ARMY WHITE SANDS MISSILE RANGE
WHITE SANDS MISSILE RANGE, NEW MEXICO

81 1 07 014

TECHNICAL REPORT

NO. 73

JANUARY 1980

ROBUST FILTERING AND SMOOTHING VIA GAUSSIAN MIXTURES

MATHEMATICAL SERVICES BRANCH
DATA SCIENCES DIVISION
US ARMY WHITE SANDS MISSILE RANGE
WHITE SANDS MISSILE RANGE, NEW MEXICO

TECHNICAL REPORT

NO. 73

Prepared by William S Agee
WILLIAM S. AGEE
Mathematician, Math Svcs Br

BARBARA A. DUNN
Mathematician, Math Svcs Br

Reviewed by Jon E. Gibson
JON E. GIBSON
Chief, Math Svcs Br

Approved by
PATRICK J. HIGGINS
Chief, Data Sciences Div

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Data Sciences Division National Range Operations Directorate White Sands Missile Range, New Mexico 88002	UNCLASSIFIED
	2b. GROUP NA AD-A093 543

3. REPORT TITLE

ROBUST FILTERING AND SMOOTHING VIA GAUSSIAN MIXTURES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

William S. Agee and Barbara A. Dunn

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
DECEMBER 1980	30	6

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO.	Technical Report No. 73
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

Robust methods provide a fresh approach to the treatment of outliers in filtering and smoothing applications. In deriving the filter and smoother equations via the conditional mean formulation or maximum a posteriori formulation the measurement noise probability density is replaced by a pseudo density which is Gaussian mixture with very heavy tails. The resulting robust filter and smoother are applied to tracking data to obtain improved estimation performance in the presence of outliers. The improvement in estimation performance is evaluated by Monte Carlo using simulated tracking data. The Monte Carlo results indicate the improvement in performance to be somewhat greater than the improvement obtained when using robust filters and smoothers derived from M-estimates.

DD FORM 1 NOV 65 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

UNCLASSIFIED
Security Classification

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Robust Estimation Filtering Smoothing						

UNCLASSIFIED
Security Classification

ABSTRACT

Robust methods provide a fresh approach to the treatment of outliers in filtering and smoothing applications. In deriving the filter and smoother equations via the conditional mean formulation or maximum a posteriori formulation the measurement noise probability density is replaced by a pseudo density which is Gaussian mixture with very heavy tails. The resulting robust filter and smoother are applied to tracking data to obtain improved estimation performance in the presence of outliers. The improvement in estimation performance is evaluated by Monte Carlo using simulated tracking data. The Monte Carlo results indicate the improvement in performance to be somewhat greater than the improvement obtained when using robust filters and smoothers derived from M-estimates.

TABLE OF CONTENTS

I. INTRODUCTION

Robust filtering and smoothing are a natural extension of the robust M-estimates of regression developed by Huber [1]. The robust M-estimates provide a natural treatment of outlying observations and have been extremely successful in dealing with outliers in other data reduction problems [2] and [3]. The extension of the M-estimates of regression to filtering and smoothing provides a fresh approach to the problems caused by outliers in filtering and smoothing applications. Robust methods for estimation have been designed to perform well when observations from contaminating distributions are present. The conventional estimation techniques of least squares, minimum variance, maximum likelihood, etc. may become useless when the observations are contaminated by outliers. In filtering and smoothing applications, outliers have often been treated by testing the filter or smoother residuals. If the residual is too large relative to some measure of dispersion of the residuals, the corresponding observation was rejected and considered to have been obtained from a contaminating distribution. Otherwise, the observation is processed normally by the filter or smoother. This procedure was often successful when only a small proportion of outliers were present in the observation sequence. Also, in order that such an outlier detection sequence be successful, a robust measure of dispersion of the filter or smoother residuals is necessary. These old methods of treating outliers in filter or smoother observations were added to the filter or smoother algorithm as an afterthought. In contrast to this the development of robust filtering and smoothing methods by the use of M-estimates provides a method of treating outlying observations which is inherent in the filter or smoother equations.

In previous reports [4] and [5] we have developed the extensions of M-estimates using a specified ψ function to the filtering and smoothing of tracking data. This application was evaluated extensively using Monte Carlo methods. The results of this evaluation showed a significant decrease in estimation error using Hampel ψ functions when small outlying observations were present. When large outliers were present, filters using M-estimates offer complete protection against the destructive effect of these outliers. Also, when no outliers were present, there was negligible loss in efficiency using filters with M-estimates as compared to using an ordinary Kalman filter.

The robust filtering and smoothing techniques developed in this report are an application of the work of Masreliez [6] on approximate non-Gaussian filtering. We assume that the filter or smoother observations are sampled from a Gaussian mixture pseudo-density. After deriving the robust filtering and smoothing equation, we give the results of an extensive Monte Carlo evaluation of these robust techniques when applied to simulated tracking data in the presence of measurement noise contaminated by outliers. We also compare these Monte Carlo results with equivalent Monte Carlo results obtained using robust filters based on Hampel ψ functions.

II. APPROXIMATE NON-GAUSSIAN FILTERING

Assume that the state of a process is given by the linear model

$$x(k+1) = \Phi(k+1,k)x(k) + u(k), \tag{1}$$

where the state vector $x(k)$ of the process is an m-vector, $u(k)$ is a Gaussian state noise vector with zero mean and covariance $Q(k)$. $\Phi(k+1,k)$ is an mxm transition matrix. Let scalar observations of the process be given by

$$Z(k) = H(k)x(k) + V(k), \tag{2}$$

where $H(k)$ is a row vector and $V(k)$ is a measurement noise error which may be contaminated by outliers.

In many situations the conditional mean provides an optimal estimate of a parameter or process. We will follow the work of Masreliez [6] in deriving an approximate conditional mean of the process specified by (1) and observed by (2). We denote the conditional mean conditioned on the observations in the set $Z^k = \{Z(1),Z(2),---,Z(k)\}$ by $\hat{x}(k|k) = E[x(k)|Z^k]$. Let $p(x(k)|Z^k)$ be the probability density of the state $x(k)$ conditioned on the observation set Z^k. Using Bayes rule $p(x(k)|Z^k)$ can be written as

$$p(x(k)|Z^k) = \frac{p(Z(k)|x(k))p(x(k)|Z^{k-1})}{p(Z(k)|Z^{k-1})} \tag{3}$$

The basic assumption in the derivation of Masreliez is that $p(x(k)|Z^{k-1})$ is Gaussian with mean $\hat{x}(k|k-1)$ and covariance $P(k|k-1)$. The conditional mean estimate of $x(k)$ is given by

$$\hat{x}(k|k) = E[x(k)|Z^k] = \int \frac{x(k)p(x(k)|Z^{k-1})p(Z(k)|x(k))dx(k)}{p(Z(k)|Z^{k-1})} \tag{4}$$

Subtracting $\hat{x}(k|k-1)$ from both side of (4)

$$\hat{x}(k|k) = \hat{x}(k|k-1) + p^{-1}(Z(k)|Z^{k-1})\int(x(k)-\hat{x}(k|k-1))p(x(k)|Z^{k-1})\,p(Z(k)|x(k))dx(k) \tag{5}$$

Using the Gaussian assumption for $p(x(k)|Z^{k-1})$, i.e., $p(x(k)|Z^{k-1}) =$

$N(x(k)-\hat{x}(k|k-1), P(k|k-1))$, we can write

$$(x(k)-\hat{x}(k|k-1))p(x(k)|Z^{k-1}) = -P(k|k-1)\frac{\partial}{\partial x(k)}p(x(k)|Z^{k-1}) \tag{6}$$

Using (6) in (5) results in

$$\hat{x}(k|k) = \hat{x}(k|k-1)-p^{-1}(Z(k)|Z^{k-1})P(k|k-1)\int\left[\frac{\partial}{\partial x(k)}p(x(k)|Z^{k-1})\right]p(Z(k)|x(k))dx(k) \tag{7}$$

Integrating (7) by parts gives,

$$\hat{x}(k|k) = \hat{x}(k|k-1) + p^{-1}(Z(k)|Z^{k-1})P(k|k-1)\int_{R^m}p(x(k)|Z^{k-1})\frac{\partial}{\partial x(k)}p(Z(k)|x(k))dx(k) \tag{8}$$

$$\text{Using}\quad \frac{\partial}{\partial x(k)}\,p(Z(k)|x(k)) = -H^T(k)\frac{\partial}{\partial Z(k)}p(Z(k)|x(k)) \tag{9}$$

in (8) we obtain

$$\hat{x}(k|k) = \hat{x}(k|k-1)-p^{-1}(Z(k)|Z^{k-1})P(k|k-1)H^T(k)\frac{\partial}{\partial Z(k)}\int_{R^m}p(Z(k)|Z^{k-1})p(Z(k)|x(k))dx \tag{10}$$

Using $p(x(k)|Z^{k-1})p(Z(k)|x(k) = p(x(k),Z(k)|Z^{k-1})$, (10) becomes

$$\hat{x}(k|k) = \hat{x}(k|k-1) + P(k|k-1)H^T(k)g(Z(k)), \tag{11}$$

where $g(Z(k))$ is the scalar

$$g(Z(k)) = p^{-1}(Z(k)|Z^{k-1})\frac{\partial}{\partial Z(k)}\, p(Z(k)|Z^{k-1}) \tag{12}$$

(11) and (12) were derived by Masreliez in [6] for approximating the minimum variance filter when the measurement noise is non-Gaussian. In order to complete the specification of this approximate filter, it is necessary to derive an expression for the conditional second moment,

$$P(k) = E\left[(x(k)-\hat{x}(k|k))(x(k)-\hat{x}(k|k))^T | Z^k \right] \tag{13}$$

An expression for computing $P(k)$ is derived similar to the derivation of $\hat{x}(k|k)$. This derivation is outlined below.

$$\begin{aligned} P(k) &= E\left[(x(k)-\hat{x}(k|k-1))(x(k)-\hat{x}(k|k-1))^T | Z^k \right] \\ &\quad - (\hat{x}(k|k)-\hat{x}(k|k-1))(\hat{x}(k|k)-\hat{x}(k|k-1))^T \end{aligned} \tag{14}$$

Let $S(k) = E\left[(x(k)-\hat{x}(k|k-1))(x(k)-\hat{x}(k|k-1))^T | Z^k \right]$. Then using (3)

$$S(k) = p^{-1}(Z(k)|Z^{k-1})\int_{R^m}(x(k)-\hat{x}(k|k-1))(x(k)-\hat{x}(k|k-1))^T p(x(k)|Z^{k-1})p(Z(k)|x(k))dx(k) \tag{15}$$

Assuming $p(x(k)|Z^{k-1})$ is Gaussian, we use (6) and integrate by parts twice to obtain.

$$S(k) = P(k|k-1) + P(k|k-1)H^T(k)\left\{ p^{-1}(Z(k)|Z^{k-1})\frac{\partial^2 p(Z(k)|Z^{k-1})}{\partial Z^2(k)}\right\}H(k)P(k|k-1) \tag{16}$$

Combining (16) with (14) and (11) gives,

$$P(k|k) = P(k|k-1) - P(k|k-1)H^T(k)G(Z(k))H(k)P(k|k-1), \tag{17}$$

where

$$G(Z(k)) = \frac{\partial g(Z(k))}{\partial Z(k)} \tag{18}$$

III. THE ROBUST CONDITIONAL MEAN FILTER

We apply (11), (12), (17), and (18) to derive a robust filter. We assume that $p(Z(k)|x(k))$ is the Gaussian mixture,

$$p(Z(k)|x(k)) = \sum_i \alpha_i N(Z(k)-H(k)x(k)-a_k^{(i)}, R_k) \tag{19}$$

In (19)

$$N(Z(k)-H(k)x(k)-a_k^{(i)}, R_k) = \left(1\Big/\sqrt{2\pi R_k}\right) \mathrm{EXP}\left\{-(Z(k)-H(k)x(k) - a_k^{(i)})^2\Big/2R_k\right\} \tag{20}$$

We do not require that $\Sigma\alpha_i = 1$ so that (19) may not be a density function, but rather a pseudo-density. Also, the sum in (19) may be infinite. Thus, we have individual Gaussians centered at $a_k^{(i)}$, each having standard deviation $\sqrt{R_k}$. The locations $a_k^{(i)}$ and the amplitudes α_i are considered design parameters of the robust filter. We obtain $p(Z(k)|Z^{k-1})$ from

$$p(Z(k)|Z^{k-1}) = \int_{R_m} p(x(k)|Z^{k-1})p(Z(k)|x(k))dx(k) \tag{21}$$

Using (19) and the Gaussian assumption for $p(x(k)|Z^{k-1})$ the convolution in (21) gives $p(Z(k)|Z^{k-1}) = \sum_i \alpha_i p(Z(k)-H(k)\hat{x}(k|k-1)-a_k^{(i)}, M(k))$ (22)

where $M(k)$ is the covariance of the residuals,

$$M(k) = H(k)P(k|k-1)H^T(k) + R_k \tag{23}$$

Using (22) in (11) and (12) we obtain the filter equation,

$$\hat{x}(k|k) = \hat{x}(k|k-1) + P(k|k-1)H^T(k)M^{-1}(k)(Z(k)-H(k)\hat{x}(k|k-1)-\bar{a}_k), \tag{24}$$

where

$$a_k = \sum_i W_i a_k^{(i)} \tag{25}$$

The weights W_i are given by

$$W_i = \frac{\alpha_i N(Z(k)-H(k)\hat{x}(k|k-1)-a_k^{(i)},M(k))}{\sum_j \alpha_j N(Z(k)-H(k)\hat{x}(k|k-1)-a_k^{(j)},M(k))} \tag{26}$$

In computing the weighted average in (25) it is not necessary to compute all the terms in the sum, since many of the weights will be zero for all practical purposes. We only need to compute those terms in (25) and the corresponding weights in (26) for which $|Z(k) - H(k)x(k|k-1)-a_k^{(i)}| \leq 3\sqrt{M(k)}$. This considerably simplifies the computation of the robust filter. The covariance of the residuals, $M(k)$, is estimated from past predicted residuals using the robust MAD estimate,

$$\sqrt{M(k)} = \underset{j=0,N-1}{\text{median}} \left| Z(k-j)-H(k-j)\hat{x}(k-j|k-j-1) \right| \Big/ .6745 \tag{27}$$

The conditional covariance, $P(k)$, is obtained using (22) in (17) and (18),

$$P(k) = P(k|k-1)-P(k|k-1)H^T(k)\left(M^{-1}(k)-M^{-2}(k)\overline{(a_k^i-\bar{a}_k)^2}\right)H^T(k)P(k|k-1), \tag{28}$$

where

$$\overline{(a_k^{(i)}-\bar{a}_k)^2} = \sum_i W_i \, (a_k^{(i)}-\bar{a}_k)^2 \tag{29}$$

Only those terms in (29) for which $|Z(k) - H(k)\hat{x}(k|k-1) - a_k^{(i)}| \leq 3\sqrt{M(k)}$ are computed.

The locations, $a_k^{(i)}$, produce a smooth pseudo-density if they are chosen at zero and odd integral multiples of $\sqrt{M(k)}$, i.e., $a_k^{(0)} = 0$, $a_k^{(i)} = \text{sgn}(i)(2|i|-1)\sqrt{M(k)}$, $|i| \geq 1$. We have also tested the filter with

$a_k^{(i)} = i \cdot \sqrt{M(k)}$, $|i| \geq 0$. Several different choices of the amplitude have been tested. The most extensive testing has been done with $\alpha_i = 1$ and $\alpha_i = 1/(|i| + 1)$.

IV. MONTE CARLO EVALUATION OF THE ROBUST FILTER

Evaluation of the robust filtering method described here has been done with a view toward eventual application to trajectory estimation. Emphasis in the evaluation is on the use of simulated rather than real trajectory data. This allows a quantitative determination of any advantages in the use of robust filtering in the presence of outliers and also any loss in efficiency using robust methods when no outliers are present. The simulated trajectory is that of a constant velocity, level flying aircraft. The filter model assumes the trajectory to have constant acceleration in three cartesian coordinates. Let x, y, z be the East, North, and Up components of trajectory position. We assume that the dynamic model for each of the coordinates is given by

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} 1 & \Delta & \Delta^2/2 \\ 0 & 1 & \Delta \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ W(k) \end{bmatrix} , \qquad (30)$$

where $\Delta = t_{k+1} - t_k$. $x_1(k)$, $x_2(k)$, $x_3(k)$ are position, velocity, and acceleration components, respectively. $W(k)$ is a zero mean Gaussian acceleration state noise with variance q. The filter observations, $Z(k)$, are scalar positions corrupted by additive noise,

$$Z(k) = H(k)x(k) + V(k),$$

with $H(k) = [1\ 0\ 0]$. The measurement noise is Gaussian with covariance $R(k)$. The measurement noise is contaminated by outliers which are generated by choosing the mean, $\mu(k)$, of the measurement noise

$$\mu(k) = \begin{cases} 0 \text{ if no outlier} \\ \mu_1(k) \text{ if outlier present} \end{cases} \qquad (31)$$

In order to decide whether or not an outlier is to be present at each time t_k, we use a two state Markov chain. Let i denote the state of the Markov chain. i = 1 is the state of no outlier present and i = 2 is the state of having an outlier present in the data. Let $P_{ij}(k)$ be the probability of a transition from state i to state j in the interval (t_{k-1}, t_k). The transition probabilities are chosen to provide a given percentage of outliers in the observations and also to generate a desired average run length of outliers. The transitions between states are realized by use of a pseudo random number generator.

The constant velocity trajectory used for evaluation is given by

$$
\begin{aligned}
x(t_{k+1}) &= x(t_k) + \dot{x}(t_{k+1}-t_k) \\
y(t_{k+1}) &= y(t_k) + \dot{y}(t_{k+1}-t_k) \\
Z(t_{k+1}) &= Z(t_k) + \dot{Z}(t_{k+1}-t_k)
\end{aligned}
\tag{32}
$$

with $\dot{x} = -550$ ft/sec, $\dot{y} = -525$ ft/sec, and $\dot{Z} = 0$. A sampling interval of $t_{k+1} = .05$ sec was used. A Monte Carlo evaluation of the filter is done by computing some statistics of the filtering errors over N filter runs. Let $\hat{x}_i(t_k)$, $\hat{y}_i(t_k)$, and $\hat{Z}_i(t_k)$ denote the filtered position estimates at time t_k for the $i^{\underline{th}}$ run and let $\tilde{x}_i(t_k) = x_i(t_k) - \hat{x}_i(t_k)$, $\tilde{y}_i(t_k) = y_i(t_k) - \hat{y}_i(t_k)$, and $\tilde{Z}_i(t_k) = Z_i(t_k) - \hat{Z}_i(t_k)$ denote the error in filtered positions for the $i^{\underline{th}}$ filter run at time t_k. Also, let $\tilde{\dot{x}}_i(t_k) = \dot{x} - \hat{\dot{x}}_i(t_k)$, $\tilde{\dot{y}}_i(t_k) = \dot{y} - \hat{\dot{y}}_i(t_k)$, and $\tilde{\dot{Z}}_i(t_k) = \dot{Z} - \hat{\dot{Z}}_i(t_k)$ denote the errors

in filtered velocities, and $\tilde{\ddot{x}}_i(t_k) = -\hat{\ddot{x}}_i(t_k)$, $\tilde{\ddot{y}}_i(t_k) = -\hat{\ddot{y}}_i(t_k)$, and $\tilde{\ddot{z}}_i(t_k) = -\hat{\ddot{z}}_i(t_k)$ be the acceleration filtering errors for the $i^{\underline{th}}$ run at time t_k. We evaluate only the RSS position, velocity, and acceleration errors,

$$
\begin{aligned}
R_i(t_k) &= \left(\tilde{x}_i^{\,2}(t_k) + \tilde{y}_i^{\,2}(t_k) + \tilde{z}_i^{\,2}(t_k)\right)^{1/2} \\
\dot{R}_i(t_k) &= \left(\tilde{\dot{x}}_i^{\,2}(t_k) + \tilde{\dot{y}}_i^{\,2}(t_k) + \tilde{\dot{z}}_i^{\,2}(t_k)\right)^{1/2} \qquad (33) \\
\ddot{R}_i(t_k) &= \left(\tilde{\ddot{x}}_i^{\,2}(t_k) + \tilde{\ddot{y}}_i^{\,2}(t_k) + \tilde{\ddot{z}}_i^{\,2}(t_k)\right)^{1/2}
\end{aligned}
$$

We compute the sample averages of the RSS error defined in (33).

$$
\begin{aligned}
\overline{R}(t_k) &= \frac{1}{N}\sum_{i=1}^{N} R_i(t_k) \\
\overline{\dot{R}}(t_k) &= \frac{1}{N}\sum_{i=1}^{N} \dot{R}_i(t_k) \qquad (34) \\
\overline{\ddot{R}}(t_k) &= \frac{1}{N}\sum_{i=1}^{N} \ddot{R}_i(t_k)
\end{aligned}
$$

In order to reduce the evaluation of the robust filter to the comparison of only a few numbers, we compute the time average of the estimation errors in (34).

$$
\begin{aligned}
\overline{\overline{R}} &= \frac{1}{M}\sum_{k=1}^{M} \overline{R}(t_k) \\
\overline{\overline{\dot{R}}} &= \frac{1}{M}\sum_{k=1}^{M} \overline{\dot{R}}(t_k) \qquad (35) \\
\overline{\overline{\ddot{R}}} &= \frac{1}{M}\sum_{k=1}^{M} \overline{\ddot{R}}(t_k)
\end{aligned}
$$

where M is the total number of filtered time points. Unless otherwise specified, M = 25, $\sqrt{R(k)}$ = 20 ft, and $P_{12}(k)$ = .05, and $P_{21}(k)$ = .5. These transition probabilities for the Markov chain provide an outlier contamination length of three. A state noise variance, q = 5, was used for all filter runs. In all runs of the robust filter, the measurement noise variance R(k), was unknown to the filter. The residual variance, which is the only quantity involving R(k) and required by the filter was estimated using the MAD estimate of (27).

Figures 1 and 2 present the average RSS position and velocity filtering errors for a Gaussian mixture robust filter with observations contaminated by various magnitudes of outliers. The Gaussian mixture filters used in generating Figs 1 and 2 used magnitudes of the Gaussians, $\alpha_i = 1/(|i|+1)$. Two different Gaussian mixture filters are represented in Figs 1 and 2, one with Gaussians at all integral multiples of the residual standard deviation, S_k, and one with Gaussians at zero and odd integral multiples of S_k. Each component of the Gaussian mixture representing $P(Z_k|Z^{k-1})$ has standard deviation, S_k.

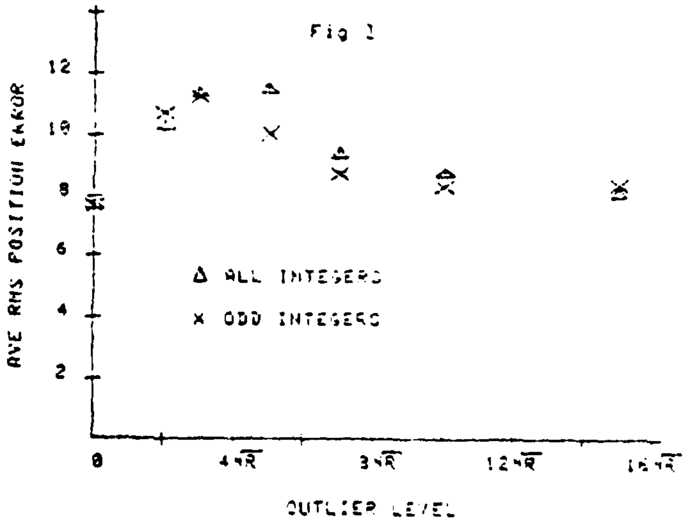
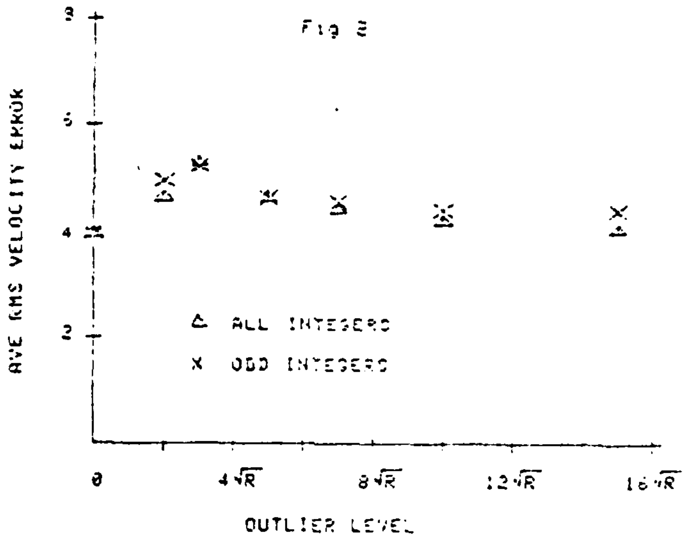
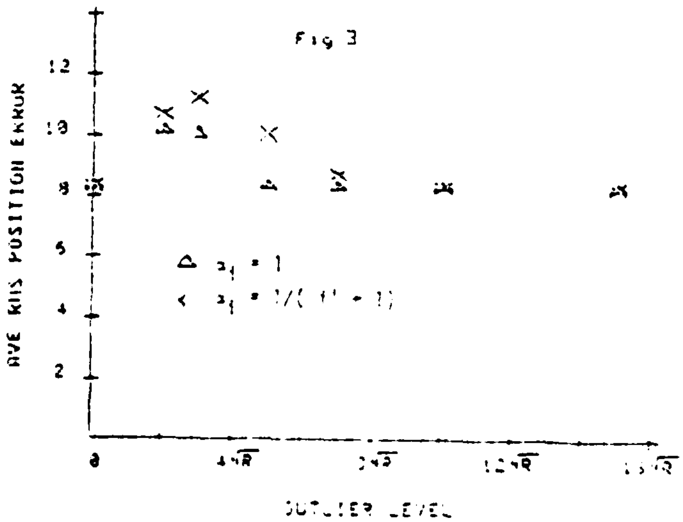


Fig 1



There is very little difference in the estimation errors of the two filters presented in Figs 1 and 2. The filter with Gaussians at only odd multiples of S_K is somewhat less computationally complex so that it might be considered the preferred filter of the two filters presented in Figs 1 and 2. Figs 3 and 4 compare the average estimation errors for two Gaussian mixture filters which have Gaussians at zero and odd integer multiples of S_K. One of these filters uses Gaussian amplitudes, $a_i = 1/(|i| + 1)$ and the other uses equal amplitudes for the Gaussians, i.e., $a_i = 1$.



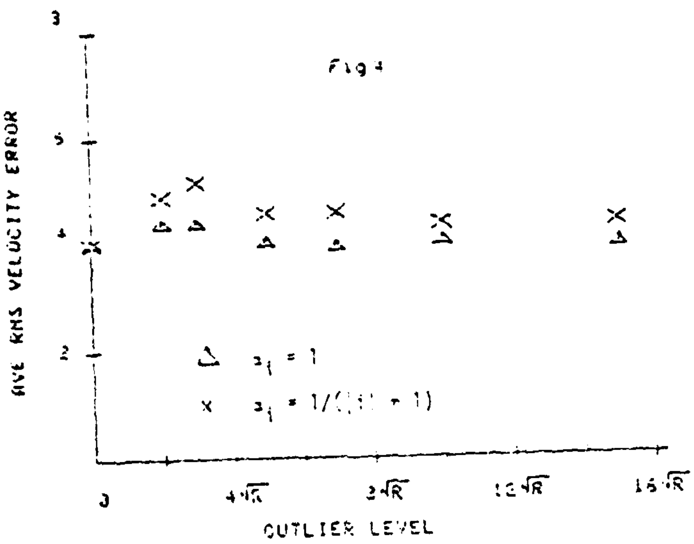


Fig 4

The robust filter which uses equal amplitudes for the Gaussians gives somewhat smaller estimation errors than the other Gaussian mixture filters evaluated. We note that using $a_i = 1$ with Gaussians placed at all integral multiples of S_k does not result in a useful filter since it has a zero influence function and therefore does not provide any error correction. Figs. 5 and 6 compare the estimation errors of the Gaussian mixture filter having $a_i = 1$ and Gaussians at odd integral multiples of S_k with the robust filter presented in [4] which uses a Hampel ψ functions with breakpoints of 1, 2, and 3 which we denote by Ha(1, 2, 3).

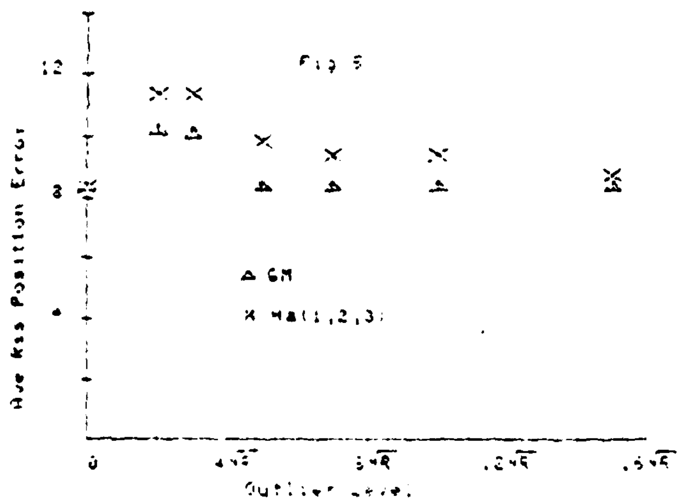
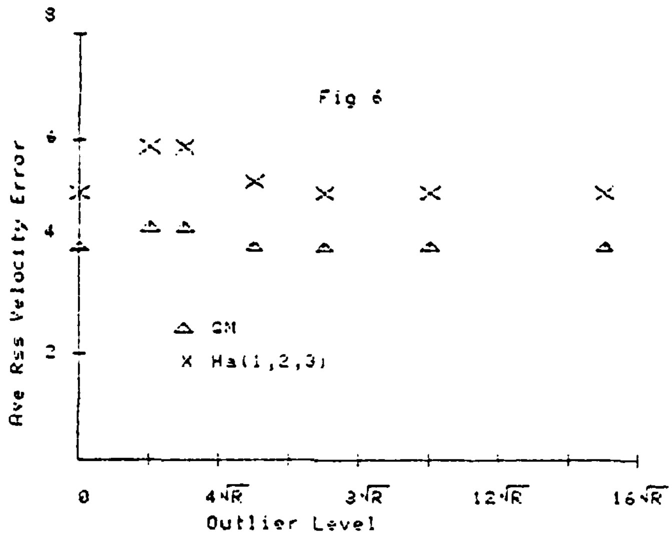


Fig 5



The Gaussian mixture filter in Figs 5 and 6 gives somewhat smaller estimation errors than the robust filter using Ha(1, 2, 3).

The robust Gaussian mixture filter was also evaluated with respect to its ability to adapt to changes in the measurement noise variance. Using the same simulated trajectory as before, the measurement noise standard deviation was taken to be $\sqrt{R(k)}$ = 20 ft for $0 \leq t \leq 25$ sec, $\sqrt{R(k)}$ = 100 ft for $25 < t \leq 40$ sec, and $\sqrt{R(k)}$ = 50 ft for $40 < t \leq 50$ sec. Figs. 7 and 8 compare the average estimation errors of the Gaussian mixture filter with $\alpha_i = 1$ and having Gaussians at zero and odd integer multiples of S_k with the estimation errors of the robust filters using the Hampel ψ functions Ha(1, 2, 3) and Ha (3, 3,3) under the above measurement noise variations. The results for Ha (1, 2, 3) and Ha (3, 3, 3) were reported in [4]. Except for the way in which the dispersion estimate, S_k, of the filter residuals is estimated, the filter using Ha (3, 3, 3) represents a conventional way of handling outliers in a Kalman filtering application. Figs. 7 and 8 show

that both the Gaussian mixture robust filter and the robust filter Ha (1, 2, 3) have considerably smaller estimation errors than the more conventional filter, Ha (3, 3, 3), for this application. Figs. 7 and 8 also indicate that the robust Gaussian mixture filter has somewhat smaller estimation errors than the robust filter using the Hampel ψ function, Ha (1, 2, 3).

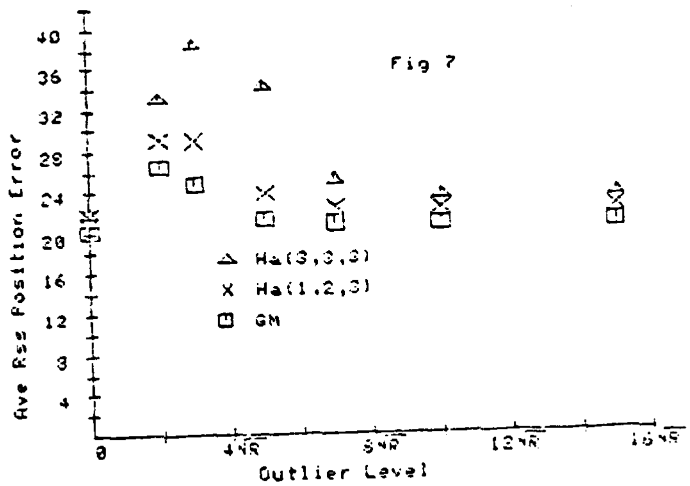


Fig 7

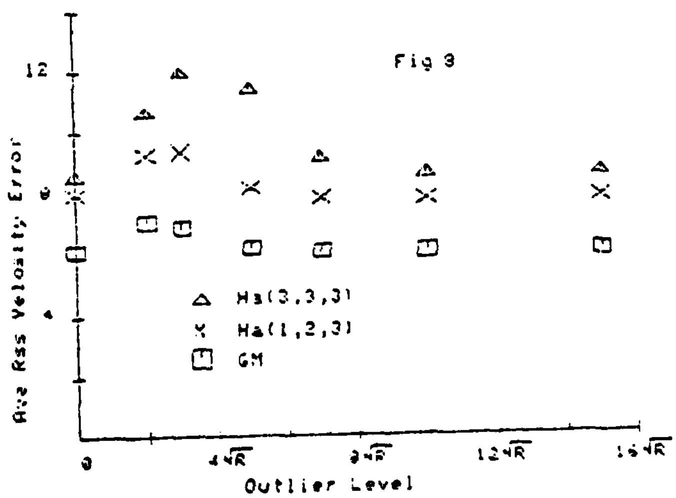


Fig 8

V. APPROXIMATE NON-GAUSSIAN SMOOTHING

In the following, some robust fixed lag smoothing equations are derived in a way similar to the derivations of the robust, Gaussian mixture filter equations, i.e., using the conditional mean derivation employed by Masreliez and assuming a Gaussian mixture pseudo density for the measurement noise. In fixed lag smoothing an estimate of the state $x(k)$ of the system described by (1) and (2) is desired using the measurements $Z(1)$, $Z(2)$, ---, $Z(k)$, ---, $Z(k+N)$. Let $\Delta Z^{k+N} = \{Z(k), Z(k+1), Z(k+2), ---, Z(k+N)\}$. Then $Z^{k+N} = Z^{k-1} \cup \Delta Z^{k+N}$. The posterior conditional density is given by

$$P(x(k)|Z^{k+N}) = \frac{P(\Delta Z^{k+N}|x(k))P(x(k)|Z^{k-1})}{P(\Delta Z^{k+N}|Z^{k-1})} \tag{36}$$

We again assume that $P(x(k)|Z^{k-1})$ is Gaussian. The conditional mean, $x(k|k+N) = E[x(k)|Z^{k+N}]$ is given by

$$\hat{x}(k|k+N) = P^{-1}(\Delta Z^{k+N}|Z^{k-1})\int_{R^m} x(k)P(\Delta Z^{k+N}|x(k))P(x(k)|Z^{k-1})dx(k) \tag{37}$$

Adding and subtracting $\hat{x}(k|k-1)$ gives

$$\hat{x}(k|k+N) = \hat{x}(k|k-1) + P^{-1}(\Delta Z^{k+N}|Z^{k-1})\int_{R^m}(x(k)-\hat{x}(k|k-1))P(x(k)|Z^{k-1}) P(\Delta Z^{k+N}|x(k))dx(k) \tag{38}$$

Assuming $P(x(k)|Z^{k-1})$ is Gaussian, we use (6) to obtain

$$\hat{x}(k|k+N) = \hat{x}(k|k-1) - P^{-1}(\Delta Z^{k+N}|Z^{k-1})P(k|k-1)\int_{R^m}\left[\frac{\partial}{\partial x(k)} P(x(k)|Z^{k-1})\right] P(\Delta Z^{k+N}|x(k))dx(k) \tag{39}$$

Integrating (39) by parts,

$$\hat{x}(k|k+N) = x(k|k-1) + P^{-1}(\Delta Z^{k+N}|Z^{k-1})P(k|k-1)\int_{R^m} P(x(k)|Z^{k-1}) \frac{\partial}{\partial x(k)} P(\Delta Z^{k+N}|x(k))dx(k) \tag{40}$$

Assuming statistical independence of the observations

$$P(\Delta Z^{k+N}|x(k)) = \prod_{j=0}^{N} P(Z(k+j)|x(k)) \tag{41}$$

Then

$$\frac{\partial}{\partial x(k)} P(\Delta Z^{k+N}|x(k)) = \sum_{j=0}^{N} Q_j \frac{\partial P(Z(k+j)|x(k))}{\partial x(k)} , \tag{42}$$

where $$Q_j = \prod_{i \neq j} P(Z(k+j|x(k)) \tag{43}$$

Assuming no state noise in the forward interval, we have

$$\frac{\partial P(Z(k+j)|x(k))}{\partial x(k)} = -\phi^T(t_{k+j}, t_k)H^T(k+j) \frac{\partial P(Z(k+j)|x(k))}{\partial Z(k+j)} \tag{44}$$

Substituting (41) - (44) into (40) gives

$$\hat{x}(k|k+N) = \hat{x}(k|k-1)-P^{-1}(\Delta Z^{k+N}|x(k))P(k|k-1) \sum_{j=0}^{N} \phi^T(t_{k+j},t_k)H^T(k+j) \frac{\partial}{\partial Z(k+j)}\int_{R^m} P(\Delta Z^{k+N}|x(k))P(x(k)|Z^{k-1})dx(k) \tag{45}$$

But

$$P(\Delta Z^{k+N}|x(k))P(x(k)|Z^{k-1}) = P(\Delta Z^{k+N},x(k)|Z^{k-1})$$

Then

$$\hat{x}(k|k+N) = \hat{x}(k|k-1) - p^{-1}(\Delta Z^{k+N}|Z^{k-1})P(k|k-1)\sum_{j=0}^{N}\Phi^{T}(t_{k+j},t_{k}) H^{T}(k+j)\frac{\partial}{\partial Z(k+j)}p(\Delta Z^{k+N}|Z^{k-1}) \tag{46}$$

In order to use (46) for robust smoothing, it is necessary to specify $p(\Delta Z^{k+N}|Z^{k-1})$ so that it has heavy tails and can be used in (46) to arrive at a useful robust smoother without too much complexity. $p(\Delta Z^{k+N}|Z^{k-1})$ can either be specified directly or computed via integration from a specification of $p(\Delta Z^{k+N}|x(k-1))$. In the latter approach the integration is too difficult when trying to achieve robustness by specifying that $p(\Delta Z^{k+N}|x(k-1)) = \prod_{j} p(Z(k+j)|x(k-1))$ and assuming that $p(Z(k+j)|x(k-1))$ is a Gaussian mixture. In the former approach assuming $p(\Delta Z^{k+N}|Z^{k-1}) = \prod_{j} p(Z(k+j)|Z^{k-1})$ with $p(Z(k+j)|Z^{k-1})$ a Gaussian mixture is easy to handle mathematically but results in poor smoothing performance. This poor performance is probably due to the fact that the variables $Z(k+j)$ conditioned on Z^{k-1} are not independent. For lack of a good method for specifying $p(\Delta Z^{k+N}|Z^{k-1})$ in (46) we abandon the conditional mean approach to robust smoothing.

A useful robust, fixed lag smoother is obtained if, instead of finding the conditional mean of $p(x(k)|Z^{k+N})$, we compute the mode of $p(x(k)|Z^{k+N})$. This estimate maximizes (36) or equivalently minimizes

$$L(x(k)) = -\log p(\Delta Z^{k+N}|x(k)) - \log p(x(k)|Z^{k-1}) \tag{47}$$

Assuming that $p(x(k)|Z^{k-1})$ is Gaussian with mean $\hat{x}(k|k-1)$ and covariance $P(k|k-1)$, (47) is equivalent to,

$$L(x(k)) = (x(k)-\hat{x}(k|k-1))^T P^{-1}(k|k-1)(x(k)-\hat{x}(k|k-1))-\log p(\Delta Z^{k+N}|x(k)) \tag{48}$$

Minimizing (48) by setting $\frac{\partial L(x(k))}{\partial x(k)} = 0$ gives

$$\hat{x}(k|k+N) = \hat{x}(k|k-1) + P(k|k-1)p^{-1}(\Delta Z^{k+N}|\hat{x}(k|k+N)) \frac{\partial p(\Delta Z^{k+N}|\hat{x}(k|k+N))}{\partial x(k|k+N)} \tag{49}$$

Assuming that

$$p(\Delta Z^{k+N}|x(k)) = \prod_{j=0}^{N} p(z(k+j)|x(k)) \tag{50}$$

Then

$$\frac{\partial p(\Delta Z^{k+N}|x(k))}{\partial x(k)} = \sum_{j=0}^{N} Q_j \frac{\partial p(Z(k+j)|x(k))}{\partial x(k)} \tag{51}$$

where

$$Q_j = \prod_{i \neq j} p(Z(k+i)|x(k)) \tag{52}$$

Also, we can write

$$\frac{\partial p(Z(k+j)|x(k))}{\partial x(k)} = -\Phi^T(t_{k+j},t_k)H^T(k+j)\ \frac{\partial p(Z(k+j)|x(k))}{\partial Z(k+j)} \tag{53}$$

Then using (50) - (53) in (49)

$$\hat{x}(k|k+N) = \hat{x}(k|k-1)-P(k|k-1)\sum_{j=0}^{N} \Phi^T(t_{k+j},t_k)H^T(k+j)p^{-1}(Z(k+j)|\hat{x}(k|k+N)) \cdot \frac{\partial p(Z(k+j)|\hat{x}(k|k+N)}{\partial Z(k+j)} \tag{54}$$

Since $\hat{x}(k|k+N)$ appears nonlinearly on the right hand side of (54), the solution of (54) will usually require iteration. We call the expression for robust smoothing given in (54) the MAP formulation of robust, fixed lag smoothing.

VI. ROBUST MAP SMOOTHING VIA GAUSSIAN MIXTURES

In order to obtain a robust, fixed lag smoother via the MAP formulation, we replace the densities, $p(Z(k+j)|x(k))$, by the Gaussian mixture pseudo-densities

$$p(Z(k+j)|x(k)) = \sum_i \alpha_i N(Z(k+j)-H(k+j)\phi(t_{k+j},t_k)x(k)-a^{(i)}_{k+j},\ R(k+j)) \tag{55}$$

In (55) we have individual Gaussians centered at $a^{(i)}_{k+j}$ having variance $R(k+j)$. The amplitude of each individual Gaussian is specified by α_i. Then

$$p^{-1}(Z(k+j)|x(k))\frac{\partial p(Z(k+j)|x(k))}{\partial Z(k+j)} = \frac{-R^{-1}(k+j)\sum_i \alpha_i(Z(k+j)-H(k+j)\phi(t_{k+j},t_k)x(k)-a^{(i)}_{k+j})N_i}{\sum_i \alpha_i N_i} \tag{56}$$

where $N_i = N(Z(k+j)-H(k+j)\phi(t_{k+j},t_k)x(k)-a^{(i)}_{k+j},R(k+j))$

(56) can be written as

$$p^{-1}(Z(k+j)|x(k))\frac{\partial p(Z(k+j)|x(k))}{\partial Z(k+j)} = -R^{-1}(k+j)(Z(k+j)-H(k+j)\phi(t_{k+j},t_k)x(k)-\bar{a}_{k+j}) \tag{57}$$

where $\bar{a}_{k+j}$ is the weighted average of the $a^{(i)}_{k+j}$,

$$\bar{a}_{k+j} = \sum_i W^{(i)}_{k+j}\ a^{(i)}_{k+j} \tag{58}$$

where the weights satisfy $\sum_i W_{k+j}^{(i)} = 1$ and

$$W_{k+j}^{(i)} = \frac{\alpha_i N_i(Z(k+j)-H(k+j)\Phi(t_{k+j},t_k)x(k)-a_{k+j}^{(i)},R(k+j))}{\sum_i \alpha_i N_i(Z(k+j)-H(k+j)\Phi(t_{k+j},t_k)x(k)-a_{k+j}^{(i)},R(k+j))} \tag{59}$$

Substituting (57) into (54) gives

$$\hat{x}(k|k+N) = \hat{x}(k|k-1)+P(k|k-1)\sum_{j=0}^{N} \Phi^T(t_{k+j},t_k)H^T(k+j)R^{-1}(k+j)$$
$$(Z(k+j)-H(k+j)\Phi(t_{k+j},t_k)\hat{x}(k|k+N)-\bar{a}_{k+j}) \tag{60}$$

Rearranging (60) results in

$$\hat{x}(k|k+N) = \hat{x}(k|k-1)+P(k)\sum_{j=0}^{N} \Phi^T(t_{k+j},t_k)H^T(k+j)R^{-1}(k+j)$$
$$(Z(k+j)-H(k+j)\Phi(t_{k+j},t_k)\hat{x}(k|k-1)-\bar{a}_{k+j}), \tag{61}$$

where

$$P^{-1}(k) = P^{-1}(k|k-1) + \sum_{j=0}^{N}\Phi^T(t_{k+j},t_k)H^T(k+j)R^{-1}(k+j)H(k+j)\Phi(t_{k+j},t_k) \tag{62}$$

Note that $\bar{a}_{k+j}$ is also dependent on $\hat{x}(k|k+N)$ so that it is necessary to iterate (61) in order to obtain a solution.

In computing the weighted averages, $\bar{a}_{k+j}$, it is not necessary to compute all the terms in the sum since many of the weights will be zero for all practical purposes. We only compute those terms for which $|Z(k+j)-H(k+j)\Phi(t_{k+j},t_k)\hat{x}^{(\alpha)}(k|k+N)-a_{k+j}^{(i)}| \leq 3\sqrt{M_{k+j}}$, where α is the iteration index and $\hat{x}^{(\alpha)}(k|k+N)$ is the approximation to $\hat{x}(k|k+N)$ at the αth interation step, $M(k+j)$ is a robust measure of the variance of the residual, $Z(k+j)-H(k+j)\Phi(t_{k+j},t_k)\hat{x}(k|k)$. This considerably simplifies the computation of the $\bar{a}_{k+j}$ since only a few of the weights, $w_{k+j}^{(i)}$, given by (59) need to be computed.

It is also necessary to have a robust estimate of the observation noise variance, R_{k+j}, to be used in the smoother in (61) and (62). We have tried several ways of obtaining a robust $R(k+j)$ and have found that the use of $R(k+j) = M(k)$ where $M(k)$ is computed from the past filtered residuals by (27) worked best in our evaluations. We form a robust measure of $M(k+j)$ by

$$M(k+j) = H(k+j)\Phi(t_{k+j},t_k)P(k|k)\Phi^T(t_{k+j},t_k) + R(k+j) \tag{63}$$

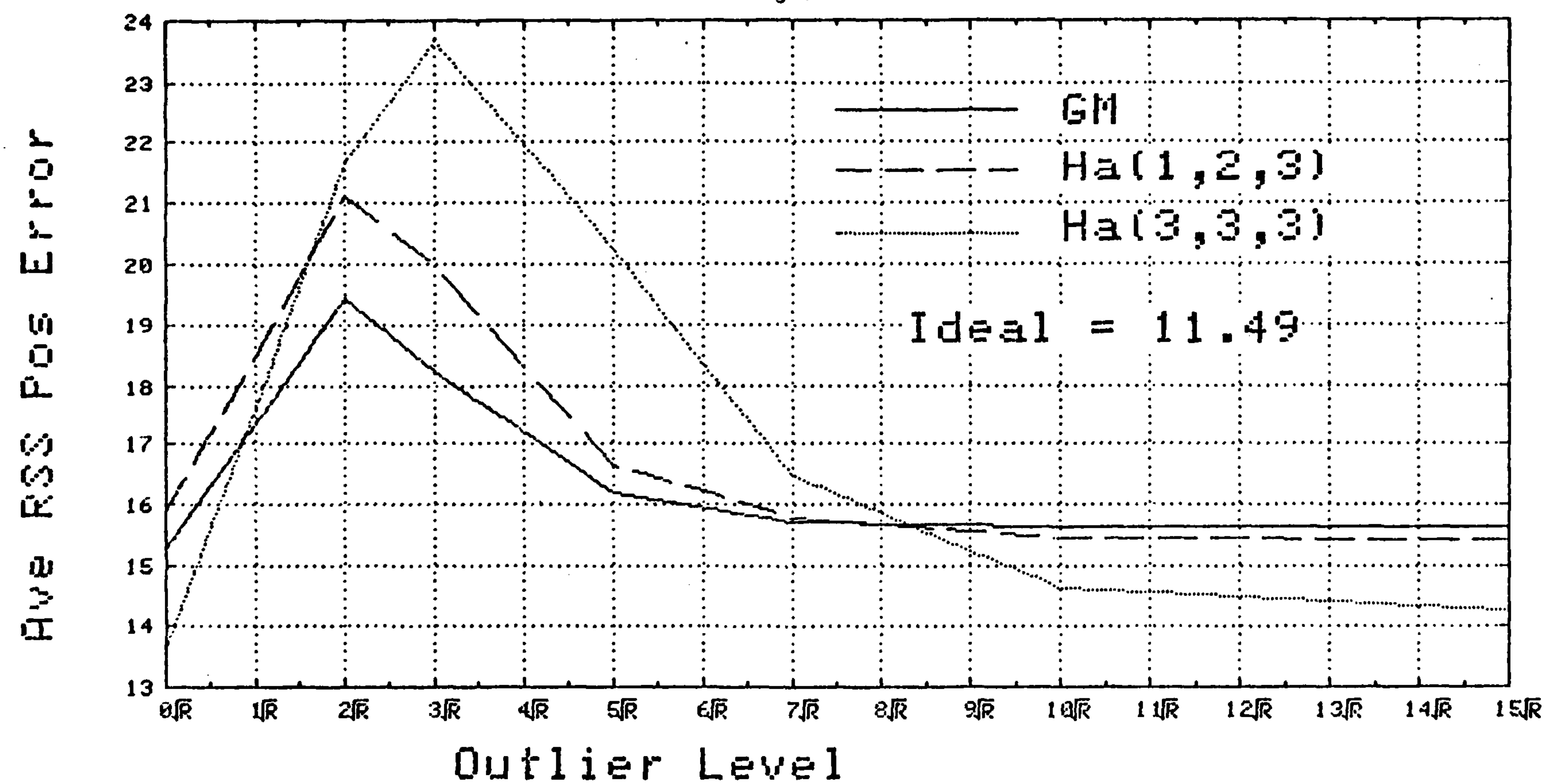
As in the robust filter we have taken equal amplitudes for each term in the Gaussian sum, i.e., $\alpha_i = 1$ and have found that taking the locations $a_{k+j}^{(i)}$ at 0 and odd integral multiples of $\sqrt{M(k+j)}$ again works well.

VII. EVALUATION OF GAUSSIAN MIXTURE ROBUST SMOOTHING

A Monte Carlo evaluation of the robust, Gaussian mixture smoother was performed using the same simulated trajectory as used for the robust filter evaluation. A smoothing interval of 1 sec. or 20 points was used. The Monte Carlo sample size for the smoother evaluation is N = 10. The measurement noise standard deviation was increased so that $\sqrt{R(k)}$ = 50 ft. The outlier contamination is the same as for the filter evaluation, i.e., a outlier contamination of 8.8% and an average outlier run length of three. The evaluation was performed using three iterations of the Gaussian mixture MAP smoother.

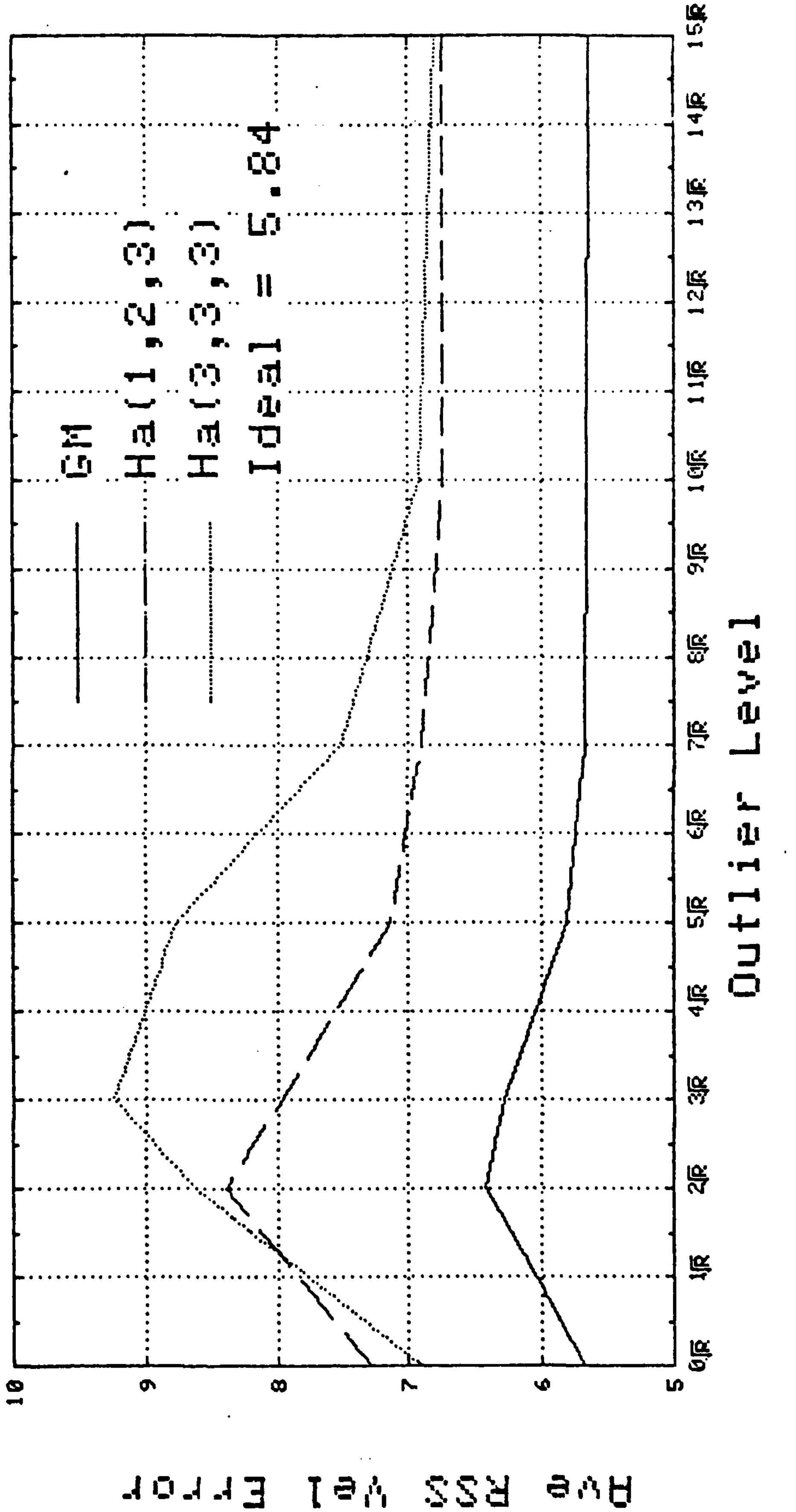
Figs 9 and 10 compare the average RSS position and velocity errors for a Gaussian mixture MAP smoother contaminated by various magnitudes of outliers with the corresponding average RSS errors of a robust MAP smoother using M-estimates with a Hampel ψ-functions which we denote by Ha(3,3,3) and Ha(1,2,3). The robust MAP smoother using Ha(3,3,3) and Ha(1,2,3) was described in [4]. Also, plotted in Figs 9 and 10 are the ideal values of the RSS errors which are obtained with an ordinary optimal smoother when no outliers are present and the noise variance is known. Figs 9 and 10 indicate that the Gaussian mixture robust smoother has some loss of efficiency (at least at the position level) when no outliers are present. Figs 9 and 10 indicate that the Gaussian mixture robust smoother has somewhat smaller estimation errors than either Ha(1,2,3) or Ha(3,3,3) in the presence of outliers.

Fig 9



24
23
22
21
20
19
18
17
16
15
14
13
Ave RSS Pos Error
0√R
1√R
2√R
3√R
4√R
5√R
6√R
7√R
8√R
9√R
10√R
11√R
12√R
13√R
14√R
15√R
Outlier Level
GM
Ha(1,2,3)
Ha(3,3,3)
Ideal = 11.49

Fig 10



GM
Ha(1,2,3)
Ha(3,3,3)
Ideal = 5.84
Outlier Level
Ave RSS Vel Error
10
9
8
7
6
5

REFERENCES

1. Huber, Peter J., "Robust Regression: Asymptotics, Conjectures and Monte Carlo", Annals of Statistics, 1, (1973), 799-821.

2. Agee, W.S. and Turner, R.H., "Application of Robust Regression to Trajectory Data Reduction," Robustness in Statistics, ed. by R.L. Launer and G. Wilkinson, Academic Press, 1979.

3. Agee, W.S. and Turner, R.H., "Robust Regression: Computational Methods for M-Estimates", Analysis and Computation Div, Tech Rpt No 66, White Sands Missile Range, 1978.

4. Agee, W.S., Turner, R.H., and Gomez, J.E., "Application of Robust Filtering and Smoothing to Tracking Data", Data Sciences Div, Tech Rpt No 71, White Sands Missile Range, 1979.

5. Agee, W.S. and Turner, R.H., "Robust Filtering and Smoothing of Tracking Data", Proc. 16th Allerton Conf. on Communication, Control, and Computing, (1979), 695-704.

6. Masreliez, Johan C., "Approximate Non-Gaussian Filtering with Linear State and Observation Relations", IEEE Trans. Auto. Conf., AC-20, (Feb 1975), 107-110.

DAT

FILM